

THESIS: MIXING TIMES OF TOURNAMENTS

SHERRY SARKAR, ADVISED BY DR. ERIC VIGODA

CONTENTS

1. Introduction	1
2. Background	3
2.1. Brief Overview of Markov Chains	3
2.2. Conductance	5
2.3. Coupling	5
3. Results	9
3.1. A Lower Bound Using Conductance for General Tournaments	9
3.2. A Markov Chain for Bipartite Tournaments	9
4. Acknowledgements	11
References	12

1. INTRODUCTION

Computer scientists have been interested in sampling uniformly from the space of graphs with a specified degree sequence. A *degree sequence* of a labelled graph, $G = (V, E)$, is a sequence of natural numbers $\{a_i\}_{i \in [n]}$ where a_i represents the degree of vertex $i \in V$. Solving this problem has several implications in networking theory. A *peer to peer network* is a decentralized data sharing network between people who dynamically join or leave the network. Random walks on the space of d regular graphs preserve optimal properties of peer to peer networks such as connectivity, low degree, and small diameter [CDG07].

In this thesis, we investigate a variant of this problem - sampling directed complete graphs of a certain out-degree sequence.

Definition 1. A *tournament* is a directed, labelled, K_n with exactly one direction assigned to each edge. The *score sequence* $\{a_n\}$ is a sequence of natural numbers such that a_i is the out-degree of vertex i .

Below, we have two examples of tournaments. The tournament on the left has a score sequence of $(2, 1, 2, 1)$. The tournament on the right has a score sequence $(6, 4, 4, 4, 3, 2, 1, 0)$.

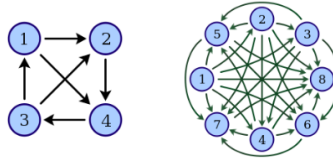


FIGURE 1. Examples of tournaments

Definition 2. A score sequence $s = (s_1, \dots, s_n)$ is *realizable* if there exists a tournament on n vertices which has s as its score sequence.

From Figure 1, we know that $(2, 1, 2, 1)$ and $(6, 4, 4, 4, 3, 2, 1, 0)$ are realizable sequences. On the other hand $(3, 2, 1, 1)$ is not a realizable sequence since the sum of the out degrees is 7 when K_4 has only 6 edges.

The first question we may ask regarding a score sequence s is whether s is realizable. It turns out there is a clean combinatorial characterization of when a score sequence is realizable.

Theorem 3. [Lan53] *Let s , without loss of generality, be a non-decreasing sequence. Then, s is realizable if and only if*

$$\sum_{i=1}^k s_i \geq \binom{k}{2}$$

for all $k \in [n]$ with equality holding for $k = n$.

The next natural question we might ask regarding a score sequence s is: how may we generate a random tournament of a given score sequence? For a practical example of this question, say we have n teams which must all play one another in a round robin tournament. For each match between teams i and j , we must decide which team's arena is used. Each team i may play exactly s_i away games. Can we sample from all valid arena assignments?

A *fully polynomial almost uniform sampler* (FPAUS) is a polynomial time algorithm which, given an input parameter s , samples elements from a set described by s , $T(s)$, in an approximately uniform manner. We rigorize this intuition with a few definitions. First, we can measure how close two distributions are using the variation distance.

Definition 4. *The **total variation distance** of a distribution μ from a distribution π over a state space Ω is*

$$d_{TV}(\mu, \pi) = \frac{1}{2} \sum_{y \in \Omega} |\mu(y) - \pi(y)|.$$

Now we can formally define a FPAUS.

Definition 5. *A **fully polynomial almost uniform sampler** (FPAUS) takes an input parameter s and an error parameter $\varepsilon > 0$ and outputs some x according to a distribution μ in time polynomial in the length of s and $\log \varepsilon^{-1}$ such that*

$$d_{TV}(\mu, \pi) \leq \varepsilon$$

where π is the uniform distribution over the desired set $T(s)$ described by s .

Markov chains are a common technique used to create a FPAUS. Markov chains are random walks on a set which, after enough steps, outputs an element according to a desired distribution on the set. The only catch is that we must determine how many steps are enough to get ε close to our desired distribution. If the number of steps (i.e the *mixing time*) is polynomial in the input size, then we call our Markov chain *rapid mixing*. We discuss technical definitions and properties of Markov chains in Section 2.

Kannan, Tetali, and Vempala [KTV99] were the first to investigate how to sample tournaments from a given score sequence. They were able to show that for regular score sequences (all vertices have the same out-degree), a simple Markov chain mixes in a polynomial of $|V|$ (but with a large and unspecified exponent) number of steps. They were able to extend this result to near regular score sequences, i.e no two out-degrees differ by more than $O(|V|^{0.75})$. In a remarkably succinct proof discovered soon after, McShine [McS00] showed that the same Markov chain can actually sample from tournaments of any score sequence, with the polynomial bound $O(n^3 \log n)$ on the mixing time. Whether McShine's bound on the mixing times of tournaments is tight is an open question. In this work, we resolve this question up to log factors. We also investigate related questions, including how to sample from orientations of a complete bipartite graph with a specified score sequence.

2. BACKGROUND

2.1. Brief Overview of Markov Chains. A Markov chain is a random process which moves from its current state to the next state based on a distribution dependent only on the current state. More formally,

Definition 6. A *Markov chain*, \mathcal{M} , is a stochastic process X_0, X_1, \dots, X_t where each X_i is a random variable over state space Ω such that

$$\mathbb{P}(X_t | X_{t-1}, \dots, X_0) = \mathbb{P}(X_t | X_{t-1}) = \mathbb{P}(X_2 | X_1).$$

The first equality comes from the Markov assumption (the next state depends only on the current) and the second inequality comes from the time invariance assumption.

There are a few ways to interpret a Markov chain. A Markov chain can be succinctly represented as a **transition matrix** P where

$$P_{ij} = \mathbb{P}(i \rightarrow j) = \mathbb{P}(X_t = j \mid X_{t-1} = i).$$

Note that the rows of P sum to 1.

Remark 7. This also gives us a way to view a Markov chain as a directed weighted graph. The transition matrix has an associated directed graph - the set of vertices is the state space Ω , and state i has an edge to state j of weight P_{ij} .

Say the probability of being in a given state $x \in \Omega$ follows some distribution μ over Ω . After a single step in the Markov chain, what is the new distribution μ' over Ω ? If we represent our distribution μ as row vector, then μ' may be succinctly described with the $|\Omega| \times |\Omega|$ transition matrix P as

$$\mu' = \mu P.$$

We denote π a *stationary distribution* of \mathcal{M} if

$$\pi = \pi M.$$

As its name suggests, a stationary distribution of \mathcal{M} is a distribution over the states which does not change after a move in \mathcal{M} . Two important questions to answer regarding the stationary distribution are: (1) whether it is unique, and (2) whether the sequence $\{\mu P^i\}_{i=1}^{\infty}$ will eventually converge to the stationary distribution from any starting distribution μ . The *Perron-Frobenius* theorem is able to answer both questions. First, we cover two preliminary definitions about non-negative matrices.

Definition 8. Let P be a non-negative matrix. P is **irreducible** if for all i, j , there exists an k such that

$$(P^k)_{i,j} > 0.$$

Recall in Remark 7, we described a way to represent a transition matrix as a directed weighted graph. For a transition matrix P , let the underlying graph be G - P is irreducible iff for all vertices i, j in G , there exists a path from i to j .

Definition 9. Let P be a $m \times m$ non-negative matrix. The **period** of index i in P is denoted as

$$p(i) = \gcd\{n \mid (P^n)_{ii} > 0\}$$

If P is irreducible, every index has the same period. We say P is **aperiodic** if for all $i \in [m]$, $p(i) = 1$.

If P is irreducible but not aperiodic, then P may have a unique stationary distribution but μP^i may never approach it. A classic example is the matrix

$$P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

which is irreducible and has a unique stationary distribution $(1/2, 1/2)$. However, since the period of each state is 2, we have that

$$(1, 0)P^i = (1, 0)$$

for even i and

$$(1, 0)P^i = (0, 1)$$

for odd i .

It is easy to guarantee a Markov chain is aperiodic by making it *lazy*. Given a Markov chain \mathcal{M} , we may make a lazy Markov chain \mathcal{M}' , where with probability $1/2$, \mathcal{M}' stays at the current state and with probability $1/2$, \mathcal{M}' moves according to \mathcal{M} . In terms of transition matrices, we have that the transition matrix P' of \mathcal{M}' is

$$P' = \frac{1}{2}P + \frac{1}{2}I$$

where P is the transition matrix of \mathcal{M} .

Now, we have all the definitions necessary to introduce the Perron-Frobenius theorem, a beautiful statement about the largest eigenvalue and its corresponding eigenspace of a non-negative matrix. While Perron-Frobenius can be formulated for non-negative matrices, we state it here specifically in the context of transition matrices.

Theorem 10 (Perron-Frobenius). *Let P be a transition matrix. If P is irreducible and aperiodic, then there exists a unique stationary distribution π and*

$$\lim_{i \rightarrow \infty} \mu P^i = \pi$$

for any starting distribution μ .

This means that if a Markov chain has an irreducible and aperiodic transition matrix, then starting at any state, the Markov chain will approach the stationary distribution π after enough steps. The question of interest to theoreticians is *how fast* will the Markov chain approach its stationary distribution.

In Section 1, we defined the total variation distance between two distributions. We apply this definition to measure how far away a Markov chain is from its stationary distribution at time t .

Definition 11. *The **total variation distance** of a Markov chain with transition matrix P at time t starting with initial state x is*

$$\Delta_x(t) = \frac{1}{2} \sum_{y \in \Omega} |P^t(x, y) - \pi(y)|.$$

Now, we may rigorously talk about how quickly a Markov chain approaches its stationary distribution.

Definition 12. *The **mixing time** of a Markov chain \mathcal{M} is*

$$\tau(\mathcal{M}) = \max_{x \in \Omega} \min\{t : \Delta_x(t') \leq \frac{1}{4} \text{ for all } t' \geq t\}.$$

A Markov chain is **rapid mixing** if it can reach a distribution $\frac{1}{4}$ away from their stationary distributions in time polynomial in the input size.

The reader may refer to Levin and Peres' text on Markov chains [LP17] for more rigorous explanations. Markov chains are the crux of all approaches outlined in this thesis. In the following subsections, we will review two common techniques to show a Markov chain is rapid mixing - conductance and coupling.

2.2. Conductance. Consider a Markov chain \mathcal{M} with transition matrix P , stationary distribution π , and state space Ω . Imagine we have a set $S \subset \Omega$ which our Markov chain is in. Given that our Markov chain starts at the stationary distribution, we have that the probability of leaving S given that we start at S is

$$\begin{aligned}\mathbb{P}(S \rightarrow \bar{S}) &= \mathbb{P}(\text{leaving } S \mid \text{starting in } S) \\ &= \frac{\mathbb{P}(\text{start in } S \text{ and then leave } S)}{\mathbb{P}(\text{being in } S)} \\ &= \frac{\sum_{x \in S, y \in \bar{S}} \pi(x) P(x \rightarrow y)}{\pi(S)}\end{aligned}$$

We call this value the conductance of a set S .

Definition 13. The *conductance* of a set $S \subset \Omega$ is defined as

$$\Phi(S) := \frac{\sum_{x \in S, y \in \bar{S}} \pi(x) P(x, y)}{\pi(S)}.$$

The *conductance* of \mathcal{M} is defined as

$$\Phi := \min_{S \subset \Omega \mid \pi(S) \leq \frac{1}{2}} \Phi(S).$$

Note that if we are in the set S , then the expected number of steps to leave S is $\frac{1}{\Phi(S)}$. To reach the stationary distribution, the Markov chain must take at least enough steps to visit states that are not in S . This motivates how the conductance can lower bound the mixing time. It turns out, conductance gives an upper bound as well. The following bounds on mixing time combine Sinclair's results on conductance [Sin93] with Aldous' inequalities for Markov chains [Ald82].

Proposition 14. [Sin93] [Ald82] *Let Φ be the conductance of Markov chain \mathcal{M} . Let π be the stationary distribution. Then,*

$$\frac{1 - 2\Phi}{4\Phi} \ln(0.5)^{-1} \leq \tau(\mathcal{M}) \leq \frac{2}{\Phi^2} (\log \max_{x \in \Omega} \pi(x)^{-1} + \ln 0.25)$$

We use conductance to show that the $O(n^3 \log n)$ upper bound McShine gave on Markov chain \mathcal{M} is tight within log factors for certain score sequences.

2.3. Coupling. The next technique we review is coupling. The idea is to consider two identical but dependent copies of a Markov chain moving in the state space Ω ; the time at which they meet is an upper bound for the mixing time of the Markov chain.

Definition 15. A *coupling* of a Markov chain \mathcal{M} is a pair of Markov chains (X_t, Y_t) such that when viewed independently, the process X_t is an exact copy of \mathcal{M} and the process Y_t is an exact copy of \mathcal{M} . More formally, for all x, x', y, y' in the state space of \mathcal{M} ,

$$\mathbb{P}(X_{t+1} = x' \mid X_t = x, Y_t = y) = \mathbb{P}(X_{t+1} = x' \mid X_t = x)$$

and

$$\mathbb{P}(Y_{t+1} = y' \mid X_t = x, Y_t = y) = \mathbb{P}(Y_{t+1} = y' \mid Y_t = y).$$

Coupling usually introduces a dependency between X_t and Y_t which brings the copies closer together as time passes. In fact, if we can show that after enough steps, the probability of X_t differing from Y_t is low, then we have a bound on the mixing time.

Theorem 16 (Coupling Lemma). *Let (X_t, Y_t) be a coupling of Markov chain \mathcal{M} . If there exists a T such that for any initial states X_0, Y_0 and for all $t \geq T$,*

$$\mathbb{P}(X_t \neq Y_t) < \frac{1}{4}$$

then the mixing time $\tau(\mathcal{M})$ is bounded by T .

One way to utilize the coupling lemma is to define a metric between states of Ω , and construct a coupling such that in each step the expected distance between any two states strictly decreases. While coupling is a powerful technique, it can sometimes be difficult to show that the expected distance between any *any* two states decreases. Bubley and Dyer [BD97] introduced a method called *path coupling* which shows it is sufficient to define a coupling between just “neighboring” states.

Definition 17. A *pre-metric* w on a set Ω is a connected, weighted, undirected graph G with vertex set Ω such that for every edge $e = (x, y) \in G$ and any path P between x and y ,

$$w(e) \leq \sum_{e' \in P} w(e').$$

If $e = (x, y)$ is an edge in G , we say x and y are **adjacent**.

In other words, if e is an edge between vertices x and y , then e constitutes the shortest path between x and y . Essentially, the edge weights in G obey the triangle inequality.

Remark 18. A pre-metric on a set Ω extends to a metric d on Ω where for any two $x, y \in \Omega$, we let $d(x, y)$ be the length of the shortest path between x and y in the pre-metric.

Theorem 19 (Path Coupling [BD97]). *Let d be a pre-metric defined on Ω . Suppose, for Markov chain \mathcal{M} , we define a coupling $(X_t, Y_t) \rightarrow (X_{t+1}, Y_{t+1})$ for only adjacent states $X, Y \in \Omega$ such that*

$$\mathbb{E}(d(X_t, Y_t)) = (1 - \alpha)\mathbb{E}(d(X_{t+1}, Y_{t+1}))$$

for some $\alpha \in (0, 1]$. Then, this coupling can be extended to a coupling for all pairs of states $X, Y \in \Omega$. In addition, this coupling gives the following bound on mixing time

$$\tau(\mathcal{M}) = O(\alpha^{-1} \log D)$$

where D is the maximum distance between any two states.

McShine used a path coupling proof to show that the mixing time of Markov chain \mathcal{M} is $O(n^3 \log n)$; we include the proof here. We first start with some properties of tournaments.

The number of triangles in a tournament is a function of just the score sequence, namely

$$N_s = \binom{n}{3} - \sum_{i=1}^n \binom{s_i}{2}$$

where N_s is the number of triangles. The combinatorial argument is as follows: there are only two types of tournaments on 3 vertices, the transitive tournament (non-triangle) and the triangle. Every three vertices form a triangle or a non-triangle, and every non-triangle has a unique source vertex v . Therefore, for each vertex v , the number of non-triangles which have v as its source vertex is

$\binom{s_i}{2}$. Since there are $\binom{n}{3}$ possible choices of three vertices, we may subtract off the number of non triangles, $\sum_{i=1}^n \binom{s_i}{2}$, to get the desired formula.

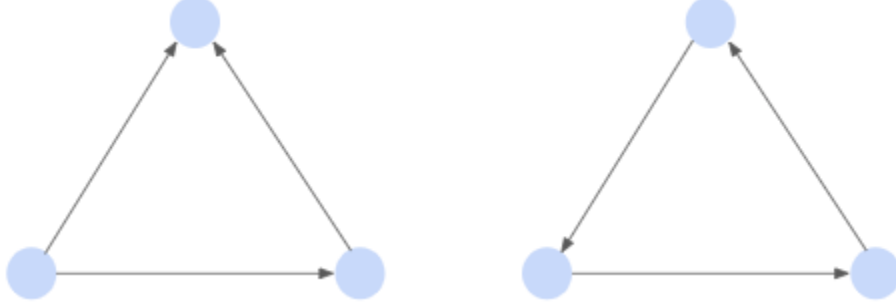


FIGURE 2. Two types of tournaments on 3 vertices

This fact motivates a very natural Markov chain, \mathcal{M} , on the tournaments of a specific score sequence s . We pick a random triangle, and we flip it.

Algorithm 1: Markov Chain for Tournaments

```

From tournament  $T$ ;
Choose a triangle  $\Delta = ABC$  u.a.r from  $T$ .
Choose  $r \in \{0, 1\}$  u.a.r.
if  $r = 0$  then
    | Flip  $\Delta$ 
else
    | Do nothing
end

```

For the initial tournament T_0 , we can enumerate all the triangles in T_0 , then choose a random triangle from this list. We then maintain a list of all the triangles in the tournament, and after a triangle flip, update the list of triangles in the new tournament. This costs $O(n)$ time, since when the edge AB is flipped, for every vertex X , ABX is either a triangle created or a triangle destroyed. Therefore, this Markov chain incurs a preprocessing cost of $O(n^3)$ with each step costing $O(n)$.

We can see that the Markov chain is aperiodic since it is lazy, so it remains to show that \mathcal{M} is irreducible.

Theorem 20. *Given any two tournaments T, T' with scores sequence s , we may transform T into T' with a sequence of no more than $\binom{n}{2} - 2$ triangle flips.*

Proof. The symmetric difference $T \oplus T'$ is the set of edges which differ between the two tournaments. Note that for each vertex v in $T \oplus T'$,

$$\text{in}(v) = \text{out}(v)$$

which means $T \oplus T'$ will be a disjoint union of directed cycles (this can be proven by induction on the edges; use BFS to find a directed cycle, remove it, and apply the inductive hypothesis).

We proceed to prove that each of the edges in a cycle C of length l in a tournament can be flipped with $l - 2$ triangle flips. In the base case of $l = 3$, C is a triangle and we are done. We assume that for a cycle of length $k \geq 3$, the statement holds and consider a cycle $C = v_1 v_2 \dots v_{k+1}$ of length $k + 1$. Consider v_1, v_2 and v_3 ; if the edge between v_1 and v_3 is directed as $v_3 v_1$, then

we have a triangle $v_1v_2v_3$. We flip this triangle and are left with the k length cycle $v_1v_3 \dots v_{k+1}$ which we know, by the inductive hypothesis, has a sequence of triangle flips which flip the cycle. If it is directed as v_1v_3 , then $v_1v_3 \dots v_{k+1}$ make a cycle of length k and we may apply the inductive hypothesis. This flips the edge between v_1 and v_3 which form the triangle $v_1v_2v_3$ which we may flip. Therefore, C may be flipped with a sequence of triangle flips; this proof also shows it may be flipped with exactly $l - 2$ triangle flips.

If $T \oplus T'$ is the made of disjoint cycles C_1, C_2, \dots, C_k , then the number of triangle flips to transform T to T' is

$$\sum_{i=1}^k l_i - 2 \leq \binom{n}{2} - 2.$$

□

Now that we know that \mathcal{M} is irreducible and aperiodic, we know it approaches its stationary distribution π . Since the number of triangles is the same for all tournaments of a given score sequence, the transition matrix is symmetric. This implies that \mathcal{M} has the uniform stationary distribution.

We next cover the path coupling argument. Our pre-metric, d (recall a pre-metric is an undirected, weighted graph satisfying the triangle inequality), on Ω is defined as: for any two tournaments $T, T' \in \Omega$ which differ by a single triangle, we have the edge (T, T') where the weight of the edge is $d(T, T') = 1$.

Consider two Markov chains X_t and Y_t which, at time t , are adjacent i.e they differ at some triangle ABC . There are three cases to consider when defining the coupling between X_t and Y_t ; let Δ_X and r_X be the random choices the Markov chain X_t uses.

- Case 1: Δ_X is exactly the triangle ABC .
- Case 2: Δ_X is a triangle which shares no edges with ABC .
- Case 3: Δ_X is a triangle adjacent to ABC .

We define X_{t+1} and Y_{t+1} based on three cases. In case 1, we assign Y_t the triangle $\Delta_Y = \Delta_X$ and the number $r_Y = 1 - r_X$. Intuitively, when the common triangle ABC is picked by X_t , Y_t picks the same triangle and does the opposite of X_t . Therefore, X_{t+1} and Y_{t+1} agree. In case 2, $\Delta_Y = \Delta_X$ and $r_Y = r_X$, i.e X_t and Y_t perform the same move. The distance between X_{t+1} and Y_{t+1} remains the same.

Case 3 is where the heart of the path coupling argument lies. Let $r_Y = r_X$. Say we pick triangle BCD in X_t . Note that in tournament Y_t , $BACD$ is a 4 cycle. If we consider the edge between D and A , if it is directed as AD , we let Δ_Y be ADB in Y_t . Then, X_{t+1} and Y_{t+1} differ at triangle ADB . If it is directed as DA , we let Δ_Y be DAC in Y_t . Then, X_{t+1} and Y_{t+1} differ at triangle DAC , and the distance between them does not change.

The distance between X_t and Y_t decreases when $\Delta_X = \Delta_Y = ABC$, the triangle differing between X_t and Y_t . The probability of picking ABC is $\frac{1}{N_s}$. Thus,

$$\mathbb{E}(d(X_{t+1}, Y_{t+1})) = (1 - \frac{1}{N_s})\mathbb{E}(d(X_t, Y_t)).$$

and we may apply the path coupling lemma to get an $O(n^3 \log n)$ mixing time.

Theorem 21. *Say s is a realizable score sequence. Then, the Markov chain described in Algorithm 1 has a mixing time of*

$$\tau(\mathcal{M}) = O(n^3 \log n)$$

3. RESULTS

3.1. A Lower Bound Using Conductance for General Tournaments. We mainly seek to answer whether the Markov chain described above produces a tight polynomial bound for sampling a random tournament uniformly. We first show a specific score sequence in which this Markov chain can mix no faster than $\Omega(n^3)$.

Theorem 22. *Let $s = (1, 1, 2, 2, d_5, d_6, \dots, d_n)$ where $d_5, d_6, \dots, d_n \geq 0$ be a score sequence such that the number of triangles in tournaments of score sequence s is $N_s = \Omega(n^3)$. The Markov chain \mathcal{M} described in Algorithm 1 cannot mix faster than $\Omega(n^3)$ on tournaments of score sequence s .*

Proof. We first prove that if the first k vertices have the property

$$\sum_{i=1}^k s_i = \binom{k}{2}$$

then the first k vertices form a strongly connected component i.e, no vertex $i \in \{1, \dots, k\}$ points to some vertex $j \in \{k+1, \dots, n\}$. Note that each edge in the induced subgraph on the first k vertices is an outwards oriented edge for some unique vertex. Therefore, the sum of the out-degrees of the first k vertices is at least $\binom{k}{2}$. Since we know that this is actually equality, there cannot exist any edges pointing out from the first k vertices into the remaining vertices.

Note that the first four out-degrees sum to 6. Therefore, the first 4 vertices form a strongly connected component. We can see from the score sequence that there are exactly two triangles in the first four vertices. Consider the set S to be all states which have the triangle on the first 3 vertices ABC . We claim that this set has low conductance. The only way to make ABC no longer a triangle is to flip a triangle adjacent to it. There is only one triangle adjacent to it since ABC is contained in a component with only 4 vertices (a 4 cycle component). For each $x \in S$, there is a $\frac{1}{N_s}$ of leaving S . Therefore,

$$\begin{aligned} \Phi(S) &= \frac{\sum_{x \in S, y \in \bar{S}} \pi(x) P(x, y)}{\pi(S)} \\ &= \frac{\sum_{x \in S, y \in \bar{S}} \frac{1}{|\Omega|} P(x, y)}{\frac{|S|}{|\Omega|}} \\ &= \frac{\sum_{x \in S, y \in \bar{S}} P(x, y)}{|S|} \\ &= \frac{|S| \frac{1}{N_s}}{|S|} = \frac{1}{N_s} \end{aligned}$$

We may conclude that for this specific score sequence, there is a lower bound of $\Omega(n^3)$ for the mixing time. \square

Remark 23. *This poses the question of whether there exists faster mixing times for certain score sequences. These conductance lower bounds on the mixing time only applies to not strongly connected tournaments with a small $O(1)$ component. One might consider whether strongly connected tournaments or near regular score sequences could possibly mix faster.*

3.2. A Markov Chain for Bipartite Tournaments. So far, we have studied how to sample from orientations of the complete graph; we may extend this question to sampling from orientations of other graphs, such as the complete bipartite graph.

Definition 24. A *bipartite tournament* is a directed, labelled, $K_{n,m}$ (the complete bipartite graph with one part of n vertices and another part of m vertices) with exactly one direction assigned to each edge. We denote the score sequence as $s = (a_1, \dots, a_n, b_1, \dots, b_m)$.

Remark 25. A bipartite graph G of degree sequence $(a_1, \dots, a_n, b_1, \dots, b_m)$ can be uniquely represented by a bipartite tournament T of score sequence $(a_1, \dots, a_n, n - b_1, \dots, n - b_m)$ using an “indicator graph”: for every edge $a_i b_j$ in G , make the directed edge $a_i b_j$ in T .

Bezáková, Bhatnagar, and Vigoda [BBV07] showed that we can approximately sample from the uniform distribution over bipartite graphs of a given degree sequence in polynomial time, though their polynomial bounds were large.

In an effort to provide a faster algorithm for sampling from bipartite tournaments of a certain score sequence s , we suggest a new Markov chain. First, we introduce some terminology. We call a tuple of 4 vertices $ABCD$ a *4-cycle* in T , if AB, BC, CD, DA are all directed edges in bipartite tournament T . Similarly, we call a tuple of 3 vertices, ABC , a *2-path* if AB and BC are directed edges in T . Let $\text{adj}(ABC, T)$ be a function denoting which four cycles of bipartite tournament T are adjacent to the 2-path ABC . More formally, let

$$\text{adj}(ABC, T) = \{ABCD \mid ABCD \text{ forms a 4-cycle in } T\}.$$

We propose the following Markov chain.

Algorithm 2: Markov Chain for Bipartite Tournaments

```

From bipartite tournament  $T$ ;
Choose 3 vertices,  $A, B, C$ , at random from  $T$ ;
if  $A, B, C$  form a path then
    WLOG, let the path read  $ABC$  ;
    if  $\text{adj}(ABC, T)$  is not empty then
        Choose a four cycle,  $ABCD$ , u.a.r from  $\text{adj}(ABC, T)$ ;
        Flip  $ABCD$  with probability  $\frac{1}{2}$ ;
    else
        end
else
    do nothing
end

```

Remark 25 shows that the Markov Chain for bipartite tournaments of a certain score sequence can be reformulated in terms of bipartite graphs of certain degree sequences.

We set out to prove that this Markov chain is ergodic and approaches the uniform distribution.

Theorem 26. *The proposed Markov chain is irreducible.*

Proof. Let T and T' be two complete bipartite tournaments of a fixed score sequence, s . Consider the symmetric difference of T and T' , which consists of edges in T with different directions from those in T' . The difference is comprised of edge disjoint even length cycles. We will show inductively that there exists a 4-cycle which can be flipped in T so that $T \ominus T'$ consists of strictly less edges. Consider the base case, where T and T' differ by a six cycle $ABCDEF$. Note that A and D are from different parts, so there must be an edge between them. If the edge is directed DA then, $ABCD$ is a 4-cycle we can flip to leave T and T' differing by $DEFA$. If it is directed AD , then $DEFA$ is a 4-cycle which may be flipped leaving T and T' differing by $ABCD$.

Say we have a cycle of length n starting with $ABCD$. We have two cases - either DA is an edge in T or it is not. If DA is an edge, then we may flip $ABCD$ in T to get a cycle of length $n - 2$. If DA is not an edge, then AD is an edge in T which means we have a cycle of length $n - 2$ containing AD . By the inductive hypothesis, there exists a 4-cycle which may be flipped to give a strictly smaller cycle. Therefore, there exists a path of 4-cycle flips in transforming T into T' . \square

Theorem 27. *The proposed Markov chain has the uniform distribution as its stationary distribution.*

Proof. We will show that the transition matrix is symmetric, i.e, for two tournaments T and T' differing by a four cycle $ABCD$,

$$\mathbb{P}(\text{flipping } ABCD \text{ in } T) = \mathbb{P}(\text{flipping } DCBA \text{ in } T').$$

Note that to flip $ABCD$, we must choose any of the following 2-paths: ABC, BCD, CDA, DAB . Each has a probability of $\frac{1}{\binom{n}{3}}$ of being chosen. Next, we will show that

$$|\text{adj}(ABC, T)| = |\text{adj}(ADC, T')|$$

by drawing an explicit bijection between $\text{adj}(ABC, T)$ and $\text{adj}(ADC, T')$. We map the cycle $ABCD$ to $ADCB$. If $ABCX$ (where X is not D) is a cycle in T , then since CXA remains a 2-path in T' . Also note that if $ABCD$ was a cycle in T , then ADC is a 2-path in T' . Therefore, $ADCX$ forms a 4-cycle. So, the bijection from $\text{adj}(ABC, T) \rightarrow \text{adj}(ADC, T')$ is

$$ABCX \rightarrow ADCX.$$

Therefore,

$$\begin{aligned} & \mathbb{P}(\text{flipping } ABCD \text{ in } T) \\ &= \frac{1}{\binom{n}{3}} \left(\frac{1}{|\text{adj}(ABC, T)|} + \frac{1}{|\text{adj}(BCD, T)|} + \frac{1}{|\text{adj}(CDA, T)|} + \frac{1}{|\text{adj}(DAB, T)|} \right) \\ &= \frac{1}{\binom{n}{3}} \left(\frac{1}{|\text{adj}(ADC, T')|} + \frac{1}{|\text{adj}(BAD, T')|} + \frac{1}{|\text{adj}(CBA, T')|} + \frac{1}{|\text{adj}(DCB, T')|} \right) \\ &= \mathbb{P}(\text{flipping } DCBA \text{ in } T') \end{aligned}$$

\square

It would be interesting to analyze the mixing time of this chain and see whether it yields a faster sampling algorithm.

4. ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. Eric Vigoda, for introducing me to the questions investigated in this thesis and mentoring me in my undergraduate research career. I would also like to thank Dr. Prasad Tetali for the comments on my thesis and for suggesting the problem of sampling from bipartite tournaments of a given score sequence. Dr. Tetali and undergraduate Daniel Hathcock have discussed the problem in relation to counting the number of acyclic orientations of the complete multi-partite graph. Lastly, I would like to thank my friends Daniel Hathcock and Shyamal Patel for getting me into theoretical computer science and making these last four years at GT incredibly fun.

REFERENCES

- [Ald82] David J. Aldous. Some inequalities for reversible Markov chains. *J. London Math. Soc.*, pages 564–576, 1982.
- [BBV07] Ivona Bezáková, Nayantara Bhatnagar, and Eric Vigoda. Sampling binary contingency tables with a greedy start. *Random Structures & Algorithms*, 30(1-2):168–205, 2007.
- [BD97] R Bubley and M.E Dyer. Path coupling: a technique for proving rapid mixing in Markov chains. *Proceedings of the 38th IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 223–231, 1997.
- [CDG07] Colin Cooper, Martin Dyer, and Catherine Greenhill. Sampling regular graphs and a peer-to-peer network. *Combinatorics, Probability and Computing*, 16:557 – 593, 2007.
- [KTV99] Ravi Kannan, Prasad Tetali, and Santosh Vempala. Simple Markov-chain algorithms for generating bipartite graphs and tournaments. *Random Structures and Algorithms*, 14:293–308, 1999.
- [Lan53] H. G. Landau. On dominance relations and the structure of animal societies: III the condition for a score structure. *The bulletin of mathematical biophysics*, 15(2):143–148, 1953.
- [LP17] David A Levin and Yuval Peres. *Markov chains and mixing times*, volume 107. American Mathematical Soc., 2017.
- [McS00] Lisa McShine. Random sampling of labeled tournaments. *The Electronic Journal of Combinatorics*, 7, 2000.
- [Sin93] Alistair Sinclair. *Algorithms for Random Generation and Counting: A Markov Chain Approach*. 01 1993.